

Solving the chicken-and-egg problem of letter detection and fixation duration in reading

ANNIE ROY-CHARLAND AND JEAN SAINT-AUBIN
Université de Moncton, Moncton, New Brunswick, Canada

AND

MICHAEL A. LAWRENCE AND RAYMOND M. KLEIN
Dalhousie University, Halifax, Nova Scotia, Canada

When asked to detect target letters while reading a text, participants miss more letters in frequent function words than in less frequent content words. According to the truncation assumption that characterizes most models of this effect, misses occur when word-processing time is shorter than letter-processing time. Fixation durations for detections and omissions were compared with fixation durations from a baseline condition when participants were searching for a target letter embedded in different words. Although, as predicted by truncation, fixation durations were longer for detections than for omissions, fixation durations for detections were also longer than those for the same words in the baseline condition, demonstrating that longer fixation durations when targets are detected are more likely to be due to demands associated with producing a detection response than to truncation. Also, contrary to predictions from the truncation assumption, the standard deviation of fixation durations for detections was larger than that from the baseline condition.

When participants are asked to read a continuous text and to circle all instances of a target letter, they make far more omissions when the target is embedded in frequent function words than when it is embedded in less frequent content words (Corcoran, 1966). This well-replicated phenomenon, known as the *missing-letter effect*, has been extensively investigated as a window on the cognitive processes involved in reading (Healy, 1994; Koriat & Greenberg, 1994). Not surprisingly, the missing-letter effect has been at the center of a vigorous theoretical debate for the last 2 decades. Dominant models of the missing-letter effect assume the *truncation* of the target detection distribution by the word identification time (Greenberg, Healy, Koriat, & Kreiner, 2004; Healy, 1976, 1994; Saint-Aubin, Klein, & Roy-Charland, 2003). More precisely, letter detection will occur for the subset of word occurrences with the longest identification time, whereas omissions will reflect the subset of occurrences with the fastest identification. Recently, Roy-Charland, Saint-Aubin, Klein, and Lawrence (2007) tested the most recent instantiation of the truncation hypothesis, the *guidance organization* (GO) model (Greenberg et al., 2004). Although most of their results invalidated the GO model, they found, consistent with the GO model, that fixation durations were longer for words for which there was a detection of the target letter than when there was an omission. However, this raises the chicken-and-egg problem: Is fixation duration longer when there is a detection because detections are more

likely when word-processing time is long, or is fixation duration extended because producing a detection response takes time? In other words, are longer word-processing times the cause or the effect of target detections? The primary aim of this study was to solve this chicken-and-egg problem. This study also tested the validity of the truncation models of the missing-letter effect by focusing on the distribution of word-processing times.

A specific instantiation of the truncation assumption in modeling the missing-letter effect is provided by the GO model (Greenberg et al., 2004), as well as by its predecessor, the *unitization account* (Healy, 1976, 1994). More specifically, according to these models, readers process text in parallel at several levels of analysis, including features, letters, syllables, and words. It is assumed that the time required to process a given letter is unaffected by the word in which it is embedded. However, as soon as the processing of the word is completed, the processing of all lower levels, including the letters, is interrupted (truncated), even if it was not completed. On this view, word frequency and word function would be distally related to omission rate by influencing word identification speed and, consequently, the time available for letter processing. High-frequency words would be identified more rapidly as a whole at the word level, due to the higher familiarity of their visual configuration. This faster identification would leave less time available for letter processing. For their part, function words would be identified more rapidly than content

A. Roy-Charland, aroycharland@laurentian.ca

words because of the contextual constraints that help to identify sentential slots where structure-supporting function words are likely to reside. Consequently, more omissions would be observed for frequent and function words than for rare and content words, because of their faster processing, leaving less time for the identification of their constituent letters (Greenberg et al., 2004).

Both the GO model and the unitization account assume the truncation of the letter-processing distribution by completion of word identification, with faster word identification resulting in omissions and slower word identification leading to detections, and they both assume the independence of word and letter processing. Figure 1 illustrates the implications of the truncation assumption. As can be seen in Figure 1, when a target letter is identified before the processing of the target word is completed, a detection response is recorded. However, if the processing of the target word is terminated before letter processing is completed, an omission occurs. By splitting the word identification distribution on the basis of whether the target letter was detected or missed, two predictions from the truncation assumption can be tested. One prediction is that the mean word-processing time on detections should be larger than that on omissions. A related prediction is that the mean of the overall word-processing time distribution (e.g., processing time for words when participants are reading and searching for a target letter not embedded in these words) should lie between the two latter means. A second prediction is that each component distribution should have a smaller standard deviation than does the overall word identification distribution.

Because the GO model and the unitization account assume independence between letter and word identification processes, Saint-Aubin et al. (2003) tested the truncation hypothesis by investigating response latencies for the letter detection task. More precisely, if it is assumed that letter and word identification speeds are independent but that letter processing is interrupted as soon as word identification is completed, it could be hypothesized that mean response latency from the letter search task should be shorter when omission rate is higher. This would occur because a higher omission rate would be due to faster word identification, earlier truncation of the letter-processing time distribution, and therefore, shorter mean detection times. In order to measure response latencies, Saint-Aubin et al. used a rapid serial visual presentation (RSVP) procedure in which words were presented one by one on a computer screen. While reading for comprehension, participants were asked to press a response key as quickly as possible when they detected the target letter. In a series of five experiments, the results revealed the usual missing-letter effect, with more omissions for function than for content words and, among content words, more omissions for more frequent words. However, contrary to the predictions of the GO model and of the unitization account, response latencies were longer when omission rate was higher: Response latencies were longer for function than for content words, and they were longer for frequent than for rare words.

In order to circumvent the inability of the GO model and of the unitization account to explain response latency results, Saint-Aubin et al. (2003) proposed a modified version of the previous models, preserving the truncation assumption while replacing the independence assumption by interaction. More precisely, under this modification, it is assumed that letter-processing time is slower for frequent function words than for less frequent content words. With this violation of the independence assumption, a truncation model can predict the pattern of response latency results obtained by Saint-Aubin et al. (see Figure 5 in Saint-Aubin et al., 2003; see Figure 2 here). High-frequency and function words being identified more quickly, a higher omission rate would be observed for letters embedded within these words. However, the processing of letters embedded in frequent function words is assumed to be less efficient than that of less frequent content words, causing a delay in the response time distribution for possible target detections. Higher omission rates would then be associated with longer response latencies when the target letter is identified, which matches the pattern obtained by Saint-Aubin et al. This modified truncation model cannot be systematically tested using response latency distribution. In effect, as the processing time of a specific

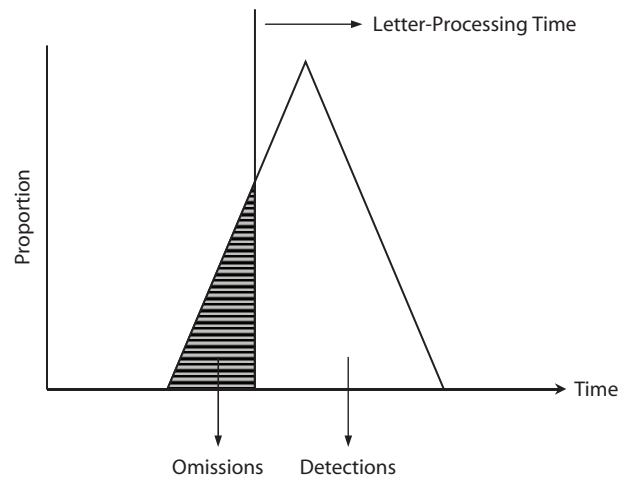


Figure 1. Relation between word- and letter-processing time and omission rate as proposed by truncation models. Word-processing time distribution is presented as a triangular distribution for graphic simplicity. The horizontal position of the vertical line represents the average time to complete the processing of a letter. Shown here as a constant for illustrative purposes, like word-processing time, letter-processing time also varies from trial to trial. Omissions are shown as the shaded region, and detections are shown as the unshaded region of the word-processing distribution. Omissions, as predicted by truncation models, are the result of word processing's being completed before letter processing is completed. Detections are the result of letter processing's being completed before word processing is completed. Word-processing times will be longer on trials in which the target is detected (unfilled portion of the distribution) than when the target is missed (shaded portion). In addition, because the whole word-processing time distribution is essentially being split into two component distributions, the variance of each component should be smaller than the variance of the whole distribution.

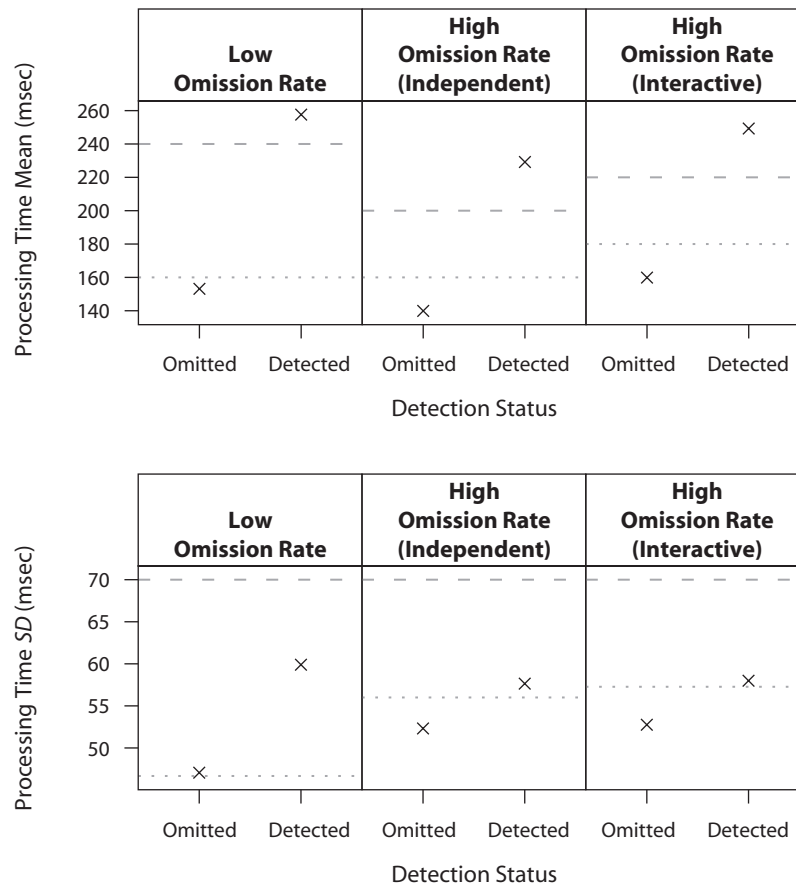


Figure 2. Simulation input and summary word-processing time (WPT) results. Lines represent input parameter values, and points represent obtained summary WPT statistics. The letter-processing time (LPT) mean is displayed as a dotted gray line in the mean panel, and standard deviation (*SD*) is displayed as a dotted gray line in the *SD* panel. Similarly, the WPT mean is displayed as a dashed gray line in the mean panel, and standard deviation is displayed as a dashed gray line in the *SD* panel. Results from simulations are plotted as points in each panel for the low omission rate, the high omission rate (independent), and the high omission rate (interactive) hypotheses.

letter would be dependent on the processing time of the word in which it is embedded, the mean letter-processing time would differ as a function of the word-processing time. Consequently, since the mean letter-processing time would no longer be constant for all words, a shorter detection time would no longer be an indicator of a shorter word-processing time (see Saint-Aubin et al., 2003). In order to test all truncation models, other processing time measures would be necessary.

The present study provided a systematic test of all models of the missing-letter effect that assume truncation of the letter-processing time distribution, regardless of whether they assume independence of letter and word processing. Word-processing speed was directly assessed by measuring viewing time as indexed by fixation durations. It must be conceded that viewing time on the target word is not a pure measure of word-processing duration, since, for instance, word processing can begin when the previous word is fixated (parafoveal preview benefit) and can extend while the next word is fixated (spillover effect).

However, although fixation duration does not provide an absolute measure of word processing, it does represent a good relative measure, because a monotonic relation is generally suggested between fixation duration and word-processing time. Consequently, fixation durations constitute the best available index of word-processing duration (Rayner, 1999). In effect, it is well established that in reading a continuous text, function words are fixated for a shorter amount of time than content words (see, e.g., Carpenter & Just, 1983) and that frequent words are also fixated for a shorter amount of time than rare words (see, e.g., Raney & Rayner, 1995; Rayner, 1998).

In a recent study, Roy-Charland et al. (2007), following Saint-Aubin and Klein (2001), monitored eye movements while participants were reading for comprehension and searching for a target letter. The results revealed the usual missing-letter effect, with more omissions for frequent function words than for less frequent content words. Most important, the results revealed that fixation durations on the target-containing word were longer when the target

letter was detected than when the target letter was omitted. These results are in line with the truncation assumption, which predicts that detections are more likely to occur when word identification time is longer. However, there is an alternative explanation for these results. It can be hypothesized that the longer fixation durations when a detection occurs are due to the planning and the production of the detection response (Griffin, 2003; Morgan & Meyer, 2005). In other words, Roy-Charland et al.'s results raise the chicken-and-egg problem: Are the longer fixations responsible for the detections, or are the detections responsible for the longer fixation durations?

The aim of the present experiment was to solve the chicken-and-egg problem by contrasting fixation durations when readers searched for a target letter embedded in the critical words with those of another group of participants asked to read exactly the same text but to search for another target letter not embedded in the critical words. The addition of a search task for a letter not embedded in the test words was necessary because, recently, Greenberg, Inhoff, and Weger (2006) found that the presence of a secondary letter search task increases gaze durations on all the words of a text, whether or not they contain a target letter.

The distribution of fixation durations in a baseline condition in which the target letter is not embedded in the critical words provides an estimate of the overall distribution of word identification times. According to the truncation hypothesis, if detection does not increase fixation duration, the fixation durations for the occurrences on which an omission or a detection occurred will be two subsets of the overall distribution from the baseline condition. Consequently, the average fixation duration for occurrences when the target letter is detected will be longer than that for the baseline condition, because it will constitute the slowest subset of the overall distribution. Furthermore, the average fixation duration for occurrences when the target letter is omitted will be shorter than that for the baseline condition, because it will constitute the fastest subset of the overall distribution. Finally, the standard deviations of fixation durations for *both* omissions and detections should be smaller than the standard deviation of the overall distribution of fixation durations, because they constitute subsets of the latter distribution.

Simulations

The above-mentioned predictions were translated in quantitative simulations using parameters derived from the present experiment. The implications of the truncation assumption for mean processing time and standard deviation of processing time are illustrated in Figure 2. Simulations were computed for instantiations of the truncation assumption, such as the GO model (Greenberg et al., 2004), that posits the independence of letter and word processing. Furthermore, simulations were also computed for a model, such as that proposed by Saint-Aubin et al. (2003), that suggests an interaction between letter and word processing. Within each panel, the summary statistics were obtained by simulating 10,000 trials of two simultaneous

processes: letter processing and word processing. On each trial, a letter-processing time and a word-processing time were obtained (by sampling from normal distributions with the means and standard deviations indicated in Figure 2), and the comparison of these determined whether the word was classified as a detection trial or an omission trial. If the trial's letter-processing time was shorter than the trial's word-processing time, the word-processing time was classified as a detection. If the trial's letter-processing time was longer than the trial's word-processing time, the word-processing time was classified as an omission. After running all 10,000 trials of the simulation, we computed the mean and standard deviation of the word-processing times from detections and omissions. The resulting data are plotted as points in each panel of Figure 2. The "Low Omission Rate" panels represent a set of input parameters for word-processing time mean, word-processing time standard deviation, and letter-processing time mean that yield a 17% omission rate. Letter-processing time standard deviation was set in a manner that matched the coefficient of variation between the word-processing time and the letter-processing time distributions. It should be noted that alternative simulations matching the letter- and word-processing time standard deviations yielded similar results. Both "High Omission Rate (Independent)" and "High Omission Rate (Interactive)" panels represent a set of input parameters that yield a 33% omission rate. Inspection of Figure 2 reveals that the "High Omission Rate (Independent)" panel differs from the "Low Omission Rate (Independent)" panel solely in the word-processing time mean parameter and that the "High Omission Rate (Interactive)" panel differs from the "Low Omission Rate" panel with respect to both word- and letter-processing time means. The patterns shown here will serve as a benchmark against which we will compare the data from the present experiment in order to test truncation models.

METHOD

Participants

Sixty-four unpaid undergraduate students (mean age of 20 years) participated in the *baseline condition* of this experiment (see the Detection of *d*: Methods and Data section for a description of the remainder of the sample).

Materials

The *des* text used by Roy-Charland et al. (2007) was used here. The text comprised 593 words. Two sets of words were selected. The first set was the *critical d words* beginning with the letter *d*, which was the target letter used in Roy-Charland et al.'s study. The second set, called the *critical p words*, comprised words beginning with the target letter *p*. For the critical *d* words, in the text, there were 24 instances of the French plural indefinite article *des* and 24 instances of three-letter control content words beginning with the letter *d*. All those words were used as critical words in Roy-Charland et al.'s study. In the present study, only a subset of these words could be used, because a requirement was implemented that none of the *d* words should be adjacent to a word in which *p* was embedded, since the latter letter was used for the search task. Consequently, the critical *d* words used here comprised 18 instances of the function word *des*, with a frequency count of 10,629 occurrences per million, and 18 instances of three-letter control content words beginning with

the letter *d*, with an average frequency count of 454 occurrences per million (e.g., *dos* [back] and *dit* [says]) (New, Pallier, Brysbaert, & Ferrand, 2004). The critical *p* words comprised 6 function words beginning with the letter *p*, with an average frequency count of 3,310 occurrences per million and an average length of four letters, and 21 content words beginning with the target letter *p*, with an average frequency count of 983 occurrences per million and an average length of six letters (New et al., 2004). In addition, the target letter *p* was embedded in 37 noncritical words. The critical *p* words were selected only to implement a search task for a letter not embedded in the critical *d* words, as a control for the additional demands of the search task on eye movement measures (Greenberg et al., 2006). Both types of critical words were never at the beginning or the end of a sentence or a line, and they were never adjacent to a punctuation mark.

Apparatus for Eye Movement Monitoring

Eye movements were measured with an SR Research Ltd. EyeLink II system. This system has high accuracy ($<0.5^\circ$) and a high sampling rate (500 Hz). The EyeLink headband has three cameras, allowing simultaneous tracking of both eyes and head position for head motion compensation. By default, only the pupil of the participant's dominant eye was tracked in our study. In the present investigation, the configurable acceleration and velocity thresholds were set to detect saccades of 0.5° or greater. A 9-point calibration procedure was used. The calibration was performed twice, and the mean deviation between both measures had to be 0.5° or smaller to be considered successful. Because, as will be seen below, between 11 and 13 lines were presented per page, this stringent calibration criterion was selected in order to reduce vertical drift problems.

Procedure

The participants took part individually in one session lasting approximately 10 min (5 min for the eyetracker calibration and 5 min for the reading task). The participants were seated about 60 cm from the screen. The participants were instructed that they would see a four-page prose passage presented on the computer screen. The text was presented in black on a white screen in Times New Roman, and it was double spaced. Between 11 and 13 lines were presented per page. At 60 cm from the screen, a horizontal character averaged 0.44° of visual angle, and the vertical spacing between lines equaled 1° of visual angle. Each page presentation was separated by a short pause that was used to verify the calibration of the system. The participants were instructed to read the text for comprehension at their normal reading rate and to search for the target letter *p*, which they were to detect by pressing the mouse key. To promote reading for comprehension, the participants were informed that they would have to answer, on a sheet of paper, three multiple-choice comprehension questions after reading the entire text. These instructions were identical to those used by Roy-Charland et al. (2007).

Detection of *d*: Methods and Data

The data set from participants reading and searching for the target letter *d* is composed of the 42 participants in Roy-Charland et al. (2007), as well as 22 participants who were added to the previous sample in order to create equal samples for the baseline condition in which the participants are searching for the target letter *p* and the *d*-search task. The overall sample of 64 participants in the *d*-search task consisted of unpaid undergraduate students with a mean age of 20 years. The eye movement measures for the 18 occurrences of the function word *des* and 18 occurrences of control content words were analyzed in order to compare them with those obtained when the participants were searching for a letter *p* not embedded in these words. Materials, apparatus, and procedure for Roy-Charland et al.'s study and the additional 22 participants added to the sample were exactly the same. Furthermore, they were also exactly the same as the ones used in the baseline condition, except that the participants were searching for *d* instead of *p*. In sum, 128 participants took part

in the study: 64 in the baseline condition and 64 in the *d*-search task condition, the latter sample composed of the 42 participants from Roy-Charland et al. and 22 further participants.

Assessment of Within-Participants Standard Deviation

In order to test the truncation hypothesis, in addition to analyses on mean fixation durations, in the present study, analyses were conducted on within-participants standard deviations. Care must be taken when calculating the standard deviation as a dependent variable; for example, when calculating the standard deviation of fixation durations on the *d* words within a given participant, there are multiple instances of each of multiple words, and simply calculating the standard deviation across all words conflates within- and between-words variability. This presents a possible confound in the circumstance in which two conditions of interest manifest different rates of word membership (e.g., comparison of the omission condition and the detection condition where there are more varied types of words in the detection condition [mostly content words] than in the omission condition [mostly the *des* function word]). Thus, the standard deviation must be calculated within participants and within words, yielding a number of standard deviation estimates that may be averaged within participants to provide a single descriptor of variability per participant.

Furthermore, it is known that the standard deviation is a biased statistical estimator such that, all other things being equal, estimates obtained from small samples will be smaller than those obtained from large samples. This bias could potentially confound the analysis of standard deviation when conditions with different sample sizes are compared (e.g., comparison of the omission condition and the detection condition, where there were fewer observations in the former than in the latter). A method by which such possible confounds may be eliminated is to equate bias across conditions by resampling. For a given set of observations for which an estimate of the standard deviation is desired, the standard deviation is calculated for two randomly chosen observations, then yet another standard deviation is calculated from yet another two randomly chosen observations, and so forth. After a large number of repetitions (we employed 10,000 in the present analyses) of drawing and calculating the standard deviation of samples of two items, a mean is calculated across the obtained standard deviation estimates, yielding a measure we dub the *mean resampled standard deviation* (MRSD). The MRSD, in effect, represents the answer to the question, What standard deviation would we observe, on average, if we drew samples of only two items from this set of observations? Thus, no matter the starting size of the set of observations, bias in the estimation of the standard deviation is equated through the use of MRSD as a measure of variability.

RESULTS

Data were scored with the EyeLink Data Viewer program, which presented the text that was read and, superimposed on it, the participants' successive landing positions. The analyses can be grouped in three categories. First, we analyzed the subset of 36 *d* words from the *d*-search task. The aim of this analysis was to establish the baseline effects with regard to fixation, omission, and response latency among the subset of words used for the main analyses, as well as reproducing the effects observed by Roy-Charland et al. (2007) with the complete data set of 64 participants. Second, the *p* words from the baseline condition were analyzed. The aim of these analyses was to document the presence of the basic missing-letter effect among the sample of words in the baseline condition. The presence of the benchmark effects with the *p* words ensures that the participants were performing in the base-

line condition as our other participants were performing the *d*-search task. In other words, this analysis provides evidence for the comparability of the two samples. Finally, the main analyses comparing fixation durations for *d* words in the *d*-search task and the baseline condition are presented. In all analyses, the .05 level of significance was adopted.

Critical *d* Words of the *d*-Search Task

In these analyses, a word was considered fixated when at least one landing position was on the word itself, be it fixated through the first pass or through a regression (see Saint-Aubin & Klein, 2001, for an application of the same criterion). Omission rates were computed by dividing the number of missed targets by the number of test words in each test word class condition (18 *des* and 18 control content words). In addition, for detected targets, response latencies were computed. Response latencies represent the time elapsed from the moment the participant's eye crossed an invisible boundary placed on the first letter of the critical words to the moment he or she pressed the response button. More precisely, during the saccade, as the eye crossed the first letter of the target words, the timer started, and the timer stopped when the buttonpress occurred. With this criterion, be the critical word fixated or not, it was possible to compute response latencies. All criteria were the same as those used by Roy-Charland et al. (2007).

Probability of fixation. Proportions of fixated target words were computed for each participant as a function of test word condition (function vs. content) by dividing the number of fixated target words by the number of occurrences. The results revealed that the probability of fixating function words (.78) was significantly smaller than the probability of fixating content words (.84) [$F(1,63) = 6.64, MS_e = 0.017$].

Omissions. Proportions of omissions were computed for each participant and each condition (function vs. content). The results revealed a significantly higher proportion of omissions when the target letter was embedded in function words (.28) than when it was embedded in content words (.16) [$F(1,63) = 24.99, MS_e = 0.017$]. Roy-Charland et al. (2007) observed an impact of fixation status on omission rates. Consequently, a second analysis was conducted while fixation status was controlled for by considering data only when words were fixated without a regression, since small cell size did not allow comparisons for words fixated with a regression and skipped occurrences. This analysis revealed the same pattern of results as that reported above.

Response latencies. Response latencies were computed by averaging response latencies for detected target letters for each participant and condition (function vs. content). The results revealed significantly longer response latencies when the target letter was embedded in function words (851 msec) than when it was embedded in content words (722 msec) [$F(1,63) = 60.54, MS_e = 8,811$]. A second analysis of response latencies for words fixated without a regression revealed the same pattern of results.

Fixation durations. All fixation duration measures used by Roy-Charland et al. (2007) were used here: (1) first-fixation duration (i.e., the duration of the first forward fixation on a target word, independent of the number of fixations made on the word), (2) single-fixation duration (i.e., the duration of the fixation when only a single forward fixation was made on the target word), (3) gaze duration (i.e., the sum of all forward fixations on a target word prior to a saccade to another word), and (4) total fixation time (i.e., the sum of all fixations, including regressions and forward fixations if there was a regression on the target word). Measures 1–3 were conditional on the target word's being fixated on the initial pass through the text, whereas Measure 4 included all fixations made on a word, regardless of whether it was made on the initial pass through or not. The analyses presented here compare fixation durations for which an omission occurred with those when a detection occurred. These analyses were expected to replicate our previous findings whose explanation raises the chicken-and-egg problem.

A first set of analyses was conducted to compare mean fixation duration measures for critical words whose targets were detected or missed. For total fixation times, 60 of the 64 participants were used, because, for 4 participants, there were no observations in one of the conditions. Similarly 59, 59, and 55 of the 64 participants were used for the first-fixation, gaze, and single-fixation durations, respectively. Inspection of the upper panel of Figure 3 reveals that fixation duration measures for omissions were shorter than those observed for detections. For single-fixation, gaze, and total fixation durations, repeated measures ANOVAs revealed significant effect of detection status (omission vs. detection) [$F(1,54) = 12.61, MS_e = 4,416$; $F(1,58) = 10.44, MS_e = 4,638$; and $F(1,59) = 23.40, MS_e = 4,742$, respectively]. For first-fixation durations, the effect was not significant [$F(1,58) = 2.60, p = .11$].

Analogous to the analyses presented above for mean fixation durations, analyses were conducted on mean resampled within-participants standard deviations for all fixation duration measures (first-fixation, single-fixation, gaze, and total). The MRSDs were computed using the technique described previously. For total fixation times, 49 of the 64 participants were used in the resampling, because, for 15 participants, there were fewer than two observations in one of the conditions, thereby precluding a measure of variability. Similarly 43, 43, and 33 of the 64 participants were used for the first-fixation, gaze, and single-fixation durations, respectively. Inspection of the lower panel of Figure 3 reveals that MRSDs for omissions were shorter than those observed for detections for all forward fixation duration measures (first-fixation, single-fixation, and gaze), but not for total fixation times. For first-fixation, single-fixation, and gaze durations, repeated measures ANOVAs revealed a significant effect of detection status (omission vs. detection) [$F(1,42) = 17.84, MS_e = 426$; $F(1,32) = 13.55, MS_e = 622$; and $F(1,42) = 12.25, MS_e = 760$, respectively]. For total fixation times, the effect of detection status did not reach significance ($F = 1.77, p = .19$). It is worth noting that a

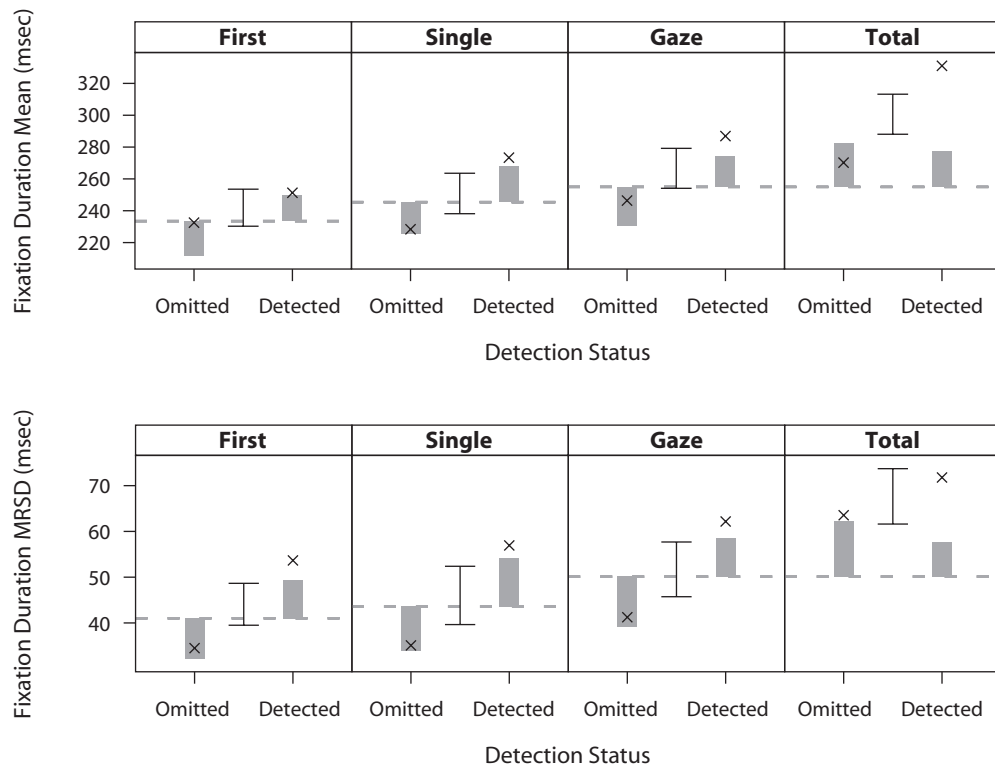


Figure 3. Means (upper panel) and mean resampled standard deviations (MRSDs) (lower panel) for the fixation duration measures (first-fixation duration, single-fixation duration, gaze duration, and total fixation time) for *d*-search omission and *d*-search detection. The horizontal dotted lines represent mean and MRSD estimates from the baseline data. Within each panel, two sets of error bars are present in the plot. The black error bar centered between *d*-search data points represents Fisher's least significant difference (FLSD); when both *d*-search data points fall outside the black bar, the two data points may be considered significantly different. The solid gray bars extending from the baseline condition dotted line represent, for each data point, the FLSD with respect to the comparison of that *d*-search point with the baseline data; when the *d*-search point falls outside the gray bar, that point may be considered significantly different from the baseline data.

similar pattern of results is obtained when using the raw within-participants standard deviations.¹

Critical *p* Words in the Baseline Condition

Probability of fixation. The results revealed that, for the critical *p* words in the baseline condition, the probability of fixating function words (.81) was significantly smaller than the probability of fixating content words (.94) [$F(1,63) = 29.32$, $MS_e = 0.017$].

Omissions. The results revealed a significantly higher proportion of omissions when the target letter was embedded in function words (.46) than when it was embedded in content words (.16) [$F(1,63) = 89.69$, $MS_e = 0.033$]. A second analysis of omission rates for words fixated without a regression revealed the same pattern of results.

Response latencies. For response latencies, 63 of the 64 participants were included in the analysis, because 1 participant made no detection on function words. The results revealed significantly longer response latencies when the target letter was embedded in function words (894 msec) than when it was embedded in content words (749 msec) [$F(1,62) = 25.16$, $MS_e = 26,352$]. A second

analysis of response latencies for words fixated without a regression revealed the same pattern of results.

Fixation durations. Gaze durations were used to assess fixation duration for the critical *p* words. Gaze duration is the sum of all forward fixations on a target word prior to a saccade to another word. For the sake of brevity, since the other estimates of fixation duration (first-fixation duration, single-fixation duration, and total fixation time) produced a pattern of results similar to that for gaze duration, only the latter will be reported here for the critical *p* words—which are not the critical data. The analysis is based on 60 of the 64 participants, because 2 participants made no omissions and 2 made no forward fixations on words for which there was an omission. The participants fixated significantly longer on *p* words when they made a detection (355 msec) than when they made an omission (262 msec) [$F(1,59) = 59.06$, $MS_e = 4,441$].

Critical *d* Words in the Baseline Condition

Probability of fixation. The aim of this analysis was to document the pattern of fixations on the 36 critical *d* words in the baseline condition. The results revealed

that the probability of fixating the function word *des* (.70) was significantly smaller than the probability of fixating the control content words (.76) [$F(1,63) = 8.34$, $MS_e = 0.014$].

Comparison of Fixation Durations on the *d* Words in the *d*-Search Task and Baseline Conditions

For all fixation duration measures, the data from the 64 participants searching for the target letter *p* were compared with the data from the 64 participants searching for the letter *d*. Data from the *d*-search task was further divided as a function of detection status—that is, occurrences in which the target letter *d* was detected and occurrences in which it was omitted. The two latter conditions will be referred to hereafter as the *d*-search detection condition and the *d*-search omission condition. The fixation durations for *d* words as a function of condition (baseline, *d*-search detection, and *d*-search omission) are presented in the upper panel of Figure 3. The three series of analyses described above establish the suitability of both data sets for conducting the main analyses of fixation durations on the critical *d* words. It is worth mentioning that the basic missing-letter effect was found among the subset of 36 *d* words from the *d*-search task sample. In addition, in the baseline condition, the participants exhibited the standard missing-letter effect with the *p* words and the standard pattern of fixation on the critical *d* words, with more skipped for *des* than for the control content words. The next series of analyses, all based on *fixation durations on the d words*, constitute the core analyses of the truncation models.

Comparison of mean fixation durations. A first set of analyses was conducted to compare fixation duration measures for the baseline condition and the *d*-search omission trials. An inspection of the upper panel of Figure 3 reveals that fixation duration measures for the *d*-search omission condition were never significantly smaller than those observed in the baseline condition. In effect, ANOVAs with condition (baseline vs. *d*-search omission) as a between-participants factor failed to reach significance for all four fixation duration measures (first-fixation, single-fixation, gaze, and total fixation durations; $F < 1$; $F = 2.84$, $p = .09$; $F < 1$; and $F = 1.23$, $p = .27$, respectively).

A second set of analyses was conducted to compare fixation duration measures for the baseline condition and the *d*-search detection trials. An inspection of the upper panel of Figure 3 reveals that all four fixation durations were significantly longer for the *d*-search detection condition than for the baseline condition. The ANOVAs with condition (baseline vs. *d*-search detection) as a between-participants factor revealed significant differences for first-fixation, single-fixation, gaze, and total fixation durations [$F(1,121) = 4.89$, $MS_e = 1,992$; $F(1,117) = 5.97$, $MS_e = 3,870$; $F(1,121) = 10.95$, $MS_e = 2,855$; and $F(1,122) = 46.44$, $MS_e = 3,856$, respectively].

Comparison of within-participants standard deviation for fixation durations. Analogous to the analyses presented above for mean fixation durations, analyses were conducted on mean resampled within-participants

standard deviations (MRSDs) for all fixation duration measures (first-fixation, single-fixation, gaze, and total fixation durations). Inspection of the lower panel of Figure 3 reveals that, for forward fixation measures (first-fixation, single-fixation, and gaze), MRSDs in the *d*-search omission condition were never significantly smaller than those observed in the baseline condition and, for total fixation time, MRSDs were significantly larger in the *d*-search omission condition than in the baseline condition. ANOVAs with condition (baseline vs. *d*-search omission) as a between-participants factor failed to reach significance for the forward fixation measures (first-fixation, single-fixation, and gaze, respectively) ($F = 1.94$, $p = .17$; $F = 2.81$, $p = .10$; and $F = 2.38$, $p = .13$, respectively). For total fixation time, MRSDs were significantly larger in the *d*-search omission condition than in the baseline condition [$F(1,111) = 4.72$, $MS_e = 1,086$].

A second series of ANOVAs was conducted for MRSDs between the baseline condition and the *d*-search detection condition. Again, the analyses were conducted with condition (baseline vs. *d*-search detection) as a between-participants factor for all fixation duration measures. Inspection of the lower panel of Figure 3 reveals that for all four fixation durations, MRSDs were significantly larger in the *d*-search detection condition than in the baseline condition. The ANOVAs revealed significant differences for first-fixation, single-fixation, gaze, and total fixation durations [$F(1,105) = 8.63$, $MS_e = 466$; $F(1,95) = 6.44$, $MS_e = 670$; $F(1,105) = 7.82$, $MS_e = 468$; and $F(1,111) = 31.20$, $MS_e = 413$, respectively]. Again, it is worth noting that a similar pattern of results was obtained when using the raw within-participants standard deviations (see note 1).

DISCUSSION

The results of the present study are clear and easy to summarize. First, both the baseline condition and the *d*-search tasks produced the usual missing-letter effect, with more omissions and longer response latencies for high-frequency function words than for lower frequency content words. It is worth noting that the present study is the third one to document response latencies, and the pattern of results is the same as that observed previously (Roy-Charland et al., 2007; Saint-Aubin et al., 2003). It is difficult to compare the size of the missing-letter effect or the accuracy level at the search task for different target letters embedded in different critical words, because omission rate varies as a function of target letter (Saint-Aubin & Poirier, 1997; Schneider & Healy, 1993), word frequency (see Healy, 1994, for a review), role (see, e.g., Koriat & Greenberg, 1994, for a review), position within the sentence (see, e.g., Moravcsik & Healy, 1998), and role predictability (Roy-Charland et al., 2007) and from occurrence to occurrence (Saint-Aubin & Klein, 2004). Nevertheless, despite all the above-mentioned factors, the results for the critical *p* words are remarkably similar to those for the critical *d* words. In addition, the results for critical *p* words in which the target letter was embedded and those for the critical *d* words in which the target letter *p* was not embedded replicate previous results in the

eye movement literature when participants were only reading (e.g., Carpenter & Just, 1983; Rayner, 1998). More precisely, the present results replicate the word class (Carpenter & Just, 1983) and word frequency effects (Raney & Rayner, 1995) observed in normal reading, with high-frequency function words being skipped more often. By replicating these previously observed patterns, the present results reveal that neither for critical words containing a target letter nor for other words does the letter search task performed in parallel to reading distort the usual pattern of eye movements (Greenberg et al., 2006; Roy-Charland et al., 2007; Saint-Aubin & Klein, 2001).

Critical Results

The results revealed that readers fixated the critical *d* words for a longer amount of time when the target letter was detected than when it was omitted. According to truncation models, this should translate into significantly longer fixation durations for detections, as compared with the baseline level provided by the baseline condition, and significantly shorter fixations for omissions, as compared with the baseline condition. The results revealed that fixations were significantly longer for detections but were never significantly shorter for omissions. It is worth noting that a similar pattern of results is obtained when function and content words are assessed separately.²

The most stringent test of truncation models was introduced by the examination of within-participants standard deviations. According to truncation models, because omissions and detections are two subsets of the same overall distribution, both should have smaller standard deviations than should the overall distribution, estimated here by the baseline condition. Contrary to truncation hypotheses, MRSDs were never significantly smaller than those in the baseline condition. In fact, with all fixation duration measures, MRSDs were significantly *larger* for detections than in the baseline condition, and they were also significantly larger for omissions with total fixation durations. According to the predictions derived from truncation hypotheses, MRSDs should not be larger than the baseline level provided by the baseline condition. As such, the present results clearly demonstrate that truncation models are wrong.

The present findings provide a clear response to the chicken-and-egg question. The fact that fixation durations for omissions are similar to the overall distribution, whereas fixation durations for detections are always slower and more variable—as evidenced by MRSDs—suggests that longer fixation durations are the consequence of detections, rather than the cause of detections. This supports the idea that longer fixation durations when a detection occurs are due to the planning and the production of the detection response. While the detection response is planned, additional time is spent gazing on the target word, which is reflected in fixation durations (Griffin, 2003; Morgan & Meyer, 2005).

Conclusion

Whether independence or an interaction between letter- and word-processing times is assumed, truncation models

make specific predictions about the mean and distribution of fixation durations (see Figure 2) that are contradicted by the present results. First, truncation models not only predict that the mean fixation duration for detections should exceed that for omissions (as was reported in Roy-Charland et al., 2007), but also predict that the overall distribution (as provided here in the baseline condition) should fall between the omission and detection means. This pattern was clearly not obtained (compare the upper panels of Figures 2 and 3). Second, truncation models predict that the within-participants standard deviations of fixation durations for trials with detected and missed targets should be smaller than the overall standard deviation. As can be seen in the lower panels of Figures 2 and 3, this pattern was not obtained either. Consequently, the GO model (Greenberg et al., 2004), the unitization account (Healy, 1994), and the interactive model proposed by Saint-Aubin et al. (2003) are not supported by the present data, because they all assume truncation of the word identification distribution. In sum, the present study opens the door to alternative models of the missing-letter effect, models that do not assume truncation of the word-processing time distribution.

AUTHOR NOTE

This research was supported by discovery grants from the Natural Sciences and Engineering Research Council of Canada to A.R.-C., to J.S.-A., and to R.M.K. M.A.L. was supported by a Julie Payette–NSERC Research Scholarship and an NSERC Canada Graduate Scholarship. This article was part of A.R.-C.'s doctoral thesis. We thank Patti Devlin for her technical assistance and Cassandra Sauvé for scoring the data. Correspondence concerning this article should be addressed to A. Roy-Charland, Department of Psychology, Laurentian University, Sudbury, ON, P3E 2C6 Canada (e-mail: aroycharland@laurentian.ca).

REFERENCES

- CARPENTER, P. A., & JUST, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275-307). New York: Academic Press.
- CORCORAN, D. W. J. (1966). An acoustic factor in letter cancellation. *Nature*, *210*, 658.
- GREENBERG, S. N., HEALY, A. F., KORIAT, A., & KREINER, H. (2004). The GO model: A reconsideration of the role of structural units in guiding and organizing text on line. *Psychonomic Bulletin & Review*, *11*, 428-433.
- GREENBERG, S. N., INHOFF, A. W., & WEGER, U. W. (2006). The impact of letter detection on eye movement patterns during reading: Reconsidering lexical analysis in connected text as a function of task. *Quarterly Journal of Experimental Psychology*, *59*, 987-995.
- GRIFFIN, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin & Review*, *10*, 603-609.
- HEALY, A. F. (1976). Detection errors on the word *the*: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception & Performance*, *2*, 235-242. doi:10.1037/0096-1523.2.2.235
- HEALY, A. F. (1994). Letter detection: A window to unitization and other cognitive processes in reading text. *Psychonomic Bulletin & Review*, *1*, 333-344.
- KORIAT, A., & GREENBERG, S. N. (1994). The extraction of phrase structure during reading: Evidence from letter detection errors. *Psychonomic Bulletin & Review*, *1*, 345-356.
- MORAVCSIK, J. E., & HEALY, A. F. (1998). Effect of syntactic role and syntactic prominence on letter detection. *Psychonomic Bulletin & Review*, *5*, 96-100. doi:10.1037/0278-7393.21.1.82

- MORGAN, J. L., & MEYER, A. S. (2005). Processing of extrafoveal objects during multiple object naming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*, 428-442. doi:10.1037/0278-7393.31.3.428
- NEW, B., PALLIER, C., BRYNSBAERT, M., & FERRAND, L. (2004). *Lexique 2: A new French lexical database*. *Behavior Research Methods, Instruments, & Computers*, *36*, 516-524.
- RANEY, G. E., & RAYNER, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology*, *49*, 151-172. doi:10.1037/1196-1961.49.2.151
- RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422. doi:10.1037/0033-2909.124.3.372
- RAYNER, K. (1999). What have we learned about eye movements during reading? In R. M. Klein & P. A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 23-56). Cambridge, MA: MIT Press.
- ROY-CHARLAND, A., SAINT-AUBIN, J., KLEIN, R. M., & LAWRENCE, M. (2007). Eye movements as direct tests of the GO model for the missing-letter effect. *Perception & Psychophysics*, *69*, 324-337.
- SAINT-AUBIN, J., & KLEIN, R. M. (2001). Influence of parafoveal processing on the missing-letter effect. *Journal of Experimental Psychology: Human Perception & Performance*, *27*, 318-334. doi:10.1037/0096-1523.27.2.318
- SAINT-AUBIN, J., & KLEIN, R. M. (2004). One missing-letter effect: Two methods of assessment. *Canadian Journal of Experimental Psychology*, *58*, 61-66. doi:10.1037/h0087440
- SAINT-AUBIN, J., KLEIN, R. M., & ROY-CHARLAND, A. (2003). Direct assessments of the processing time hypothesis for the missing-letter effect. *Journal of Experimental Psychology: Human Perception & Performance*, *29*, 1191-1210. doi:10.1037/0096-1523.29.6.1191
- SAINT-AUBIN, J., & POIRIER, M. (1997). The influence of the word function in the missing-letter effect: Further evidence from French. *Memory & Cognition*, *25*, 666-676.
- SCHNEIDER, V. I., & HEALY, A. F. (1993). Detecting phonemes and letters in text: Interactions between different types and levels of processes. *Memory & Cognition*, *21*, 739-751.

NOTES

1. Analyses were conducted on raw within-participants standard deviations for all fixation duration measures (first-fixation, single-fixation, gaze, and total fixations). Standard deviations were weighted for the number of observations per participant for each condition. A first set of analyses was conducted on within-participants standard deviations for fixation durations, with detection status (omission vs. detection) as a

within-participants factor. For first-fixation, single-fixation, gaze, and total fixation times, the results revealed that standard deviations were significantly larger for detections (109, 111, 130, and 152) than for omissions (77, 80, 95, and 118) [$F(1,49) = 15.64$, $MS_e = 1,643$; $F(1,47) = 11.42$, $MS_e = 2,023$; $F(1,49) = 12.06$, $MS_e = 2,531$; and $F(1,53) = 16.38$, $MS_e = 1,885$, respectively].

A second series of analyses were conducted to compare the within-participants standard deviations between the baseline condition and the *d*-search omission condition. ANOVAs with condition (baseline vs. *d*-search omission) as a between-participants factor were conducted for all fixation duration measures. For each of the fixation duration measures—first-fixation, single-fixation, gaze, and total fixation times—the *d*-search omission condition (77, 80, 95, and 118) did not differ significantly from the baseline condition (90, 92, 109, and 127) [$F(1,112) = 3.05$, $p = .08$; $F(1,110) = 1.99$, $p = .16$; $F(1,112) = 2.35$, $p = .13$; and $F(1,116) = 0.92$, $p = .34$, respectively].

A third series of ANOVAs was conducted between the baseline condition and the *d*-search detection condition. Again, the analyses were conducted with condition (baseline vs. *d*-search detection) as a between-participants factor for all fixation duration measures. For each of the fixation duration measures—first-fixation, single-fixation, gaze, and total fixation times—the results revealed that standard deviations were significantly larger in the *d*-search detection condition (106, 108, 129, and 157) than in the baseline condition (90, 92, 109, and 127) [$F(1,126) = 6.54$, $MS_e = 1,195$; $F(1,126) = 5.88$, $MS_e = 1,478$; $F(1,126) = 8.92$, $MS_e = 1,391$; and $F(1,126) = 17.71$, $MS_e = 1,634$, respectively].

2. An analysis was conducted to compare gaze durations on the critical *d* words for the baseline (content = 256 msec, function = 243 msec) and the *d*-search omission trials (content = 269 msec, function = 245 msec), with word class (function vs. content) as a factor. For the *d*-search omission condition, 39 of the 64 participants were used, because for 25 participants there was no observation in one of the conditions. The results revealed a marginally significant effect of word class [$F(1,101) = 3.50$, $MS_e = 4,805$, $p = .06$], but neither the effect of condition nor the interaction reached significance (both F s < 1).

A second analysis was conducted to compare gaze durations on the critical *d* words for the baseline condition (content = 256 msec, function = 243 msec) and the *d*-search detection trials (content = 286 msec, function = 273 msec), with word class (function vs. content) as a factor. The results revealed a marginally significant effect of word class [$F(1,126) = 3.80$, $MS_e = 2,874$, $p = .05$] and a significant effect of condition [$F(1,126) = 16.10$, $MS_e = 3,547$], but the interaction did not reach significance ($F < 1$).

(Manuscript received December 4, 2008;
revision accepted for publication April 6, 2009.)